BIOMEDIN 215 DATA DRIVEN MEDICINE

With the spread of electronic health records, increasingly large data repositories of clinical and other patient derived data are being built. These databases are large and difficult for any one specialist to analyze. To find the hidden associations within such data, we review methods for large-scale **data-mining on electronic medical records**, methods in natural language processing and **text-mining of medical records**, methods for using **ontologies for tagging of unstructured clinical notes**.

SCHEDULE: TUE, THU 2:15 PM - 3:30 PM
LOCATION: ALWAY BUILDING, ROOM M112
CREDITS: 3

The course has four modules. The first module will review the medical data miner's tool-kit—including the use of ontologies for data-mining. The remaining modules will review three problem areas and computational methods used in that problem area via a set of 5 lectures ending in a "mini project" as home work. Each module will cover a new application area (e.g., predicting readmission, drug safety surveillance, clinical text mining) and a new method (e.g. association rules, logistic regression).

The course will use real, de-identified, large size patient datasets (millions of patients range) that are made available for a final research project associated with the course.

This course is also offered in a 2 credit version (BIOMEDIN 225) which meets at the same time but does not require a final project.